Stamen

What We Learned: The Evolution of a Pandemic Through a Data Lens

When Johns Hopkins' Coronavirus Resource Center announced that its COVID-19 dashboard was going away, many of us in the data design & visualization world took special note. It was the end of an era.

For three years when the only certainty was uncertainty, it was the resource to see what was going on day-to-day. During that time many of us struggled to find clarity and confidence in our data: Should I make that trip to the place? Should we reopen the office? Is it getting better or worse out there? Johns Hopkins was a stable, authoritative voice that provided confidence in our understanding of the world; but there are many other stories about pandemic data, some of which are still being written today.

A Covid Data Journey in Three Acts

COVID-19 launched a thousand data visualizations. Before its appearance, people regularly consumed visualized statistics generated by news, sporting events, and

research. But most of those events produced a set of data in a predictable timeframe. The event would occur and generate data. Then the data would be presented, and it was over. With the exception of financial services and some specialized fields, ongoing day-to-day tracking of data wasn't something most people did — and certainly not something they did to make decisions about their everyday lives.

By contrast, the COVID-19 pandemic evolved daily and uncertainly. It presented a rapidly shifting landscape of data collection, interpretation, and most of all for Stamen, data storytelling and visualization.

Scatter plots, the arc of a line, the shading and scaling of a digital map — they all became the brush and canvas as we captured and mapped our understanding of this new reality.

Stamen collaborated on three unique COVID data projects, each with their own data peculiarities. In order to make sense of the pandemic in the early days, for the first project we explored the pandemic's impact by starting with immediate data sources that captured official case counts. We then had an opportunity to focus on data acquisition and dashboard reporting of on-premises pathogens for building managers. Finally, the third project reached a height of certainty lacking in the other two: We visualized wastewater data, the new benchmark for pathogen tracking, by capitalizing on rigorously tested data that provided volumes of insight.

Looking back, we see three distinct acts to the overarching narrative of our work during the pandemic:

Act 1
On-the-Fly Data Collection: Seizing the Opportunity of a Terrifying Moment

Act 2
Self-Reporting: The Challenge of Sparse and Sporadic Data

Act 3
Wastewater Data: Setting the Gold Standard

ACT 1

On-the-Fly Data Collection: Seizing the Opportunity of a Terrifying Moment

In the months just before the pandemic, we worked on an exciting project with a team from the University of California San Francisco School of Medicine in the Dean's Office of Population Health and Health Equity. Of course, as it was originally conceived, this project had nothing to do with the onset of a global pandemic.

We call this effort the UCSF Health Atlas. Intended to help researchers explore how the environment influences health at a local level in California, we developed the UCSF Health Atlas to curate publicly available data to inform an interactive map. The overall project goal was to give researchers new insights into the social determinants of health.

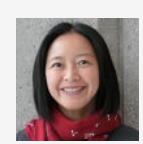
Exploration of Public Health Principles

Dr. Debby Oh, Health Atlas project lead and epidemiologist at UCSF, knew that well-designed, interactive maps have the unique ability to display the environmental, public health, and socioeconomic disparities communities experience.

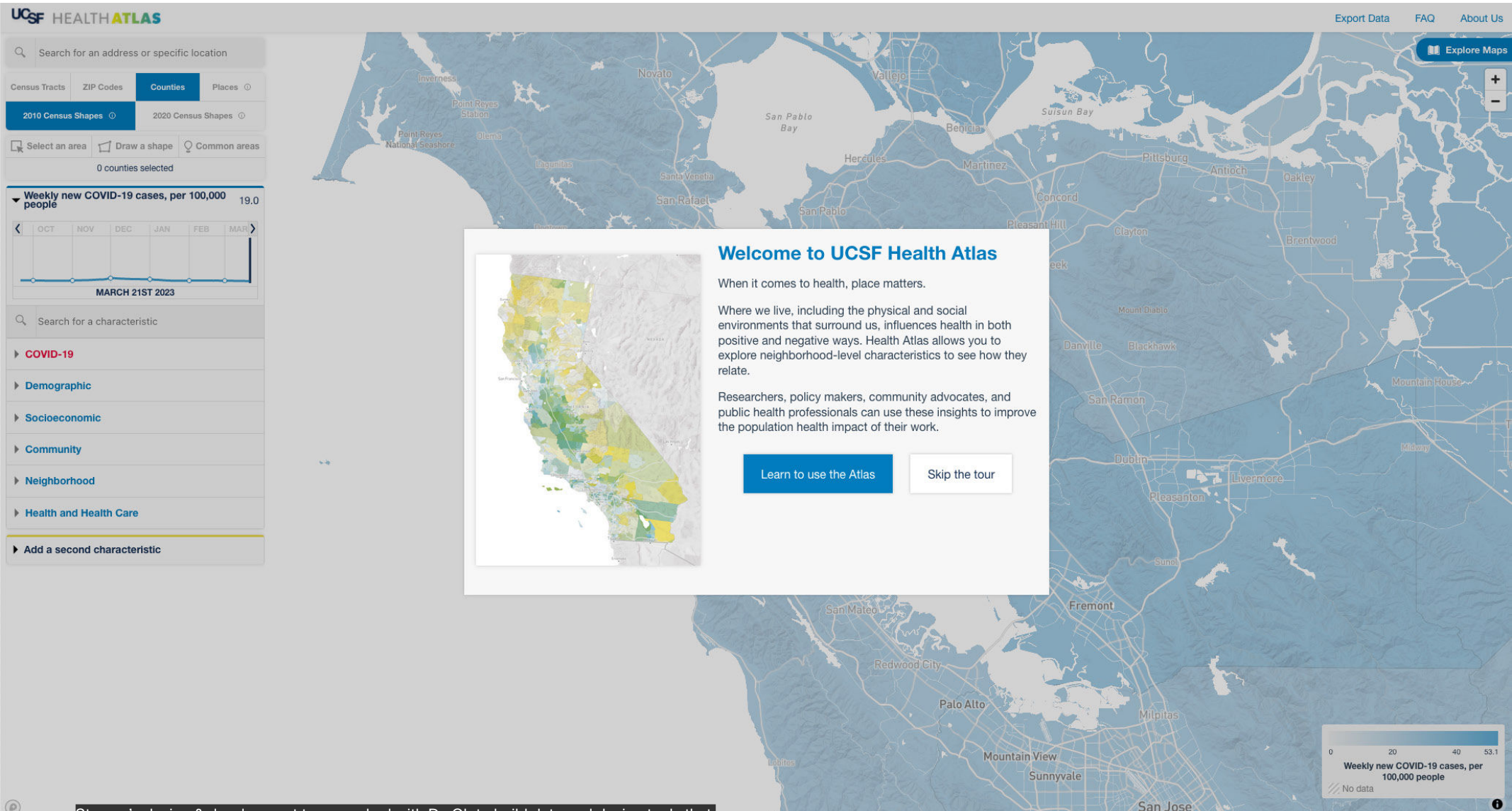
She also knew that building an end-to-end data pipeline to report data from publicly available sources, such as the American Community Census Bureau Survey, the 500 Cities Project from the Centers of Disease Control and Prevention, the US Department of Agriculture Food Access Research Atlas, and many more, required sophisticated technical solutions.

Empowering users of the Health Atlas to intuitively pool or parse that data, as well as easily download it for further analysis, required expert data design skills.

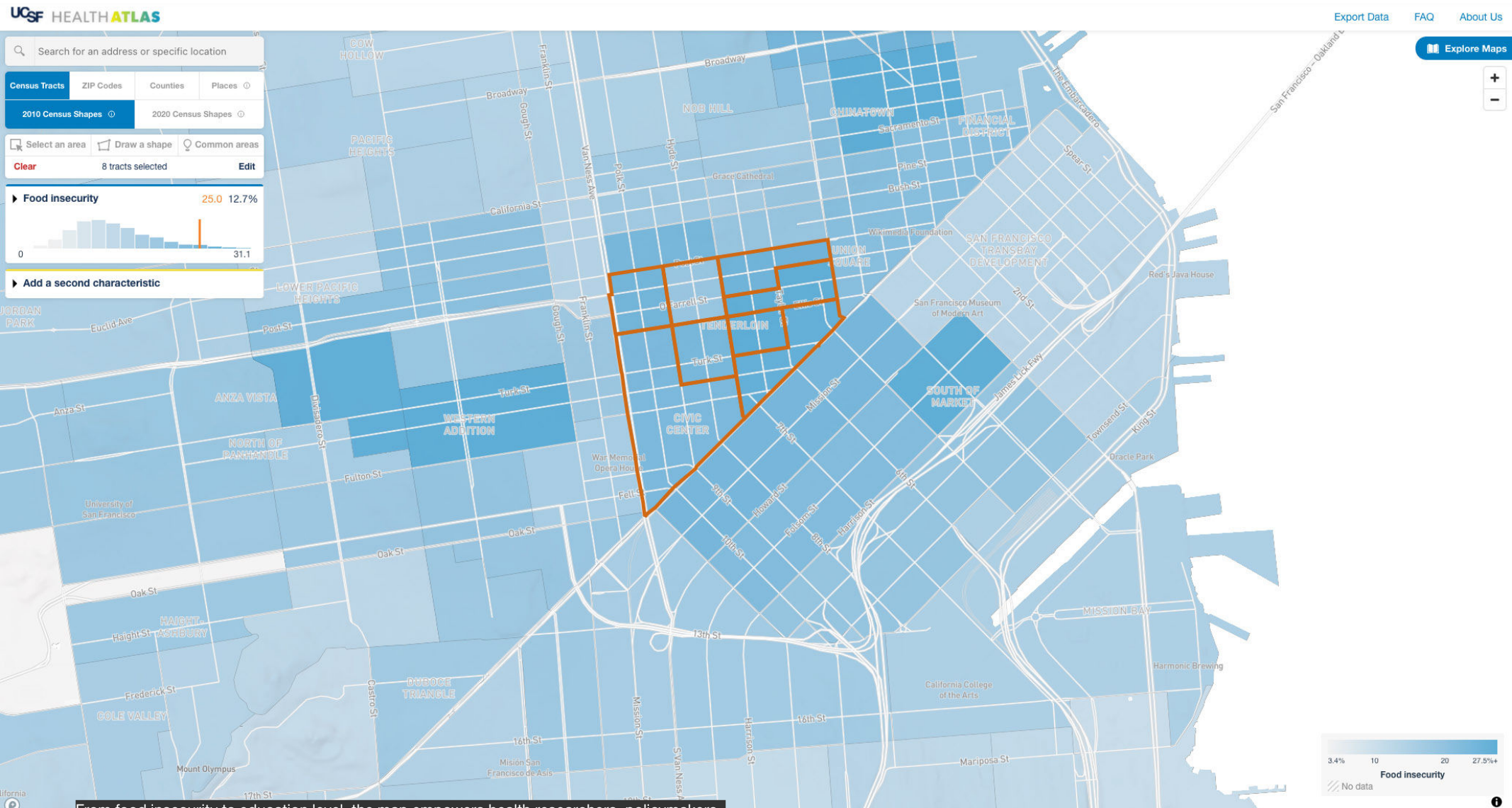
"The vision was to provide researchers an easy-to-use tool to explore what factors can impact health on a population level. We wanted to provide beautiful maps, well-informed and well-written content, the ability to pool geographies, a comparison of variables, and downloadable data."



Debby Oh
Ph.D., UCSF



Stamen's design & development team worked with Dr. Oh to build data and design tools that render data as a crisp and useful interactive map. Via an expertly hierarchized UI that fosters exploration, users can zoom in or pan out to discover more than 100 social determinants of health and health metrics across the state — all at census tract and county levels.



From food insecurity to education level, the map empowers health researchers, policymakers, and community advocates to explore insights about environmental trends and health inequities. Users can easily select and view characteristics, create custom queries to learn about specific areas, and can compare multiple health or environmental scenarios at once.

Food insecurity: UCSF Health Atlas

COVID-19 Enters the World Stage

As the UCSF Health Atlas was nearing completion, COVID-19 began its relentless march across the globe. An opportunity beckoned: to study COVID's spread and health impacts through communities by leveraging the existing data set on the Health Atlas. Together, UCSF and Stamen pivoted and ideated ways to address this by including COVID data as part of the Atlas.

Prior to the onset of the pandemic, the UCSF Heath Atlas was powered by an archived library of data sources. While the project pulled data from numerous sources, that data was static and unchanging. Showcasing the power of design & development collaboration, Stamen's designers, technologists, cartographers, and data scientists met the challenge of illuminating sets of local-view data and finding new data sources (e.g., the Los Angeles Times) that reported COVID-19 numbers in real-time.

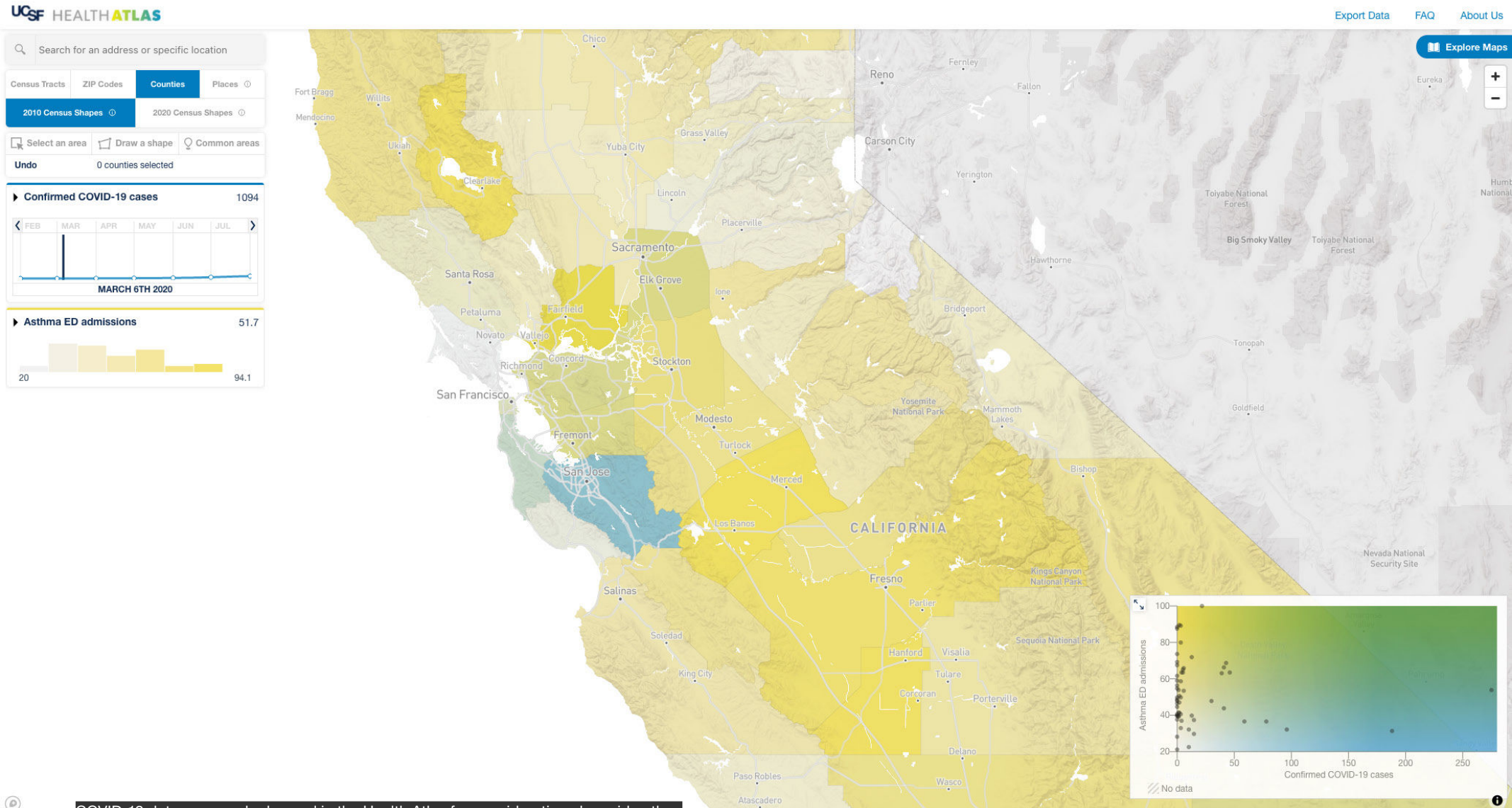
Public officials and organizations could now consult the map to know when school closures were needed and where events could safely be held.

Quick problem-solving and end-to-end technical innovation enabled Stamen to incorporate salient new use cases (image, next page) into the project that were well outside the focus of the UCSF Health Atlas's original goals.

Themes of Act 1's COVID-19 Data Project

Rapidly iterating and problem-solving to pull real-time data into the Health Atlas fundamentally changed how the site functioned, thereby increasing its public impact. Stamen homed in on a beautiful way to visualize truly on-the-fly data sets. In doing so, the Health Atlas illuminated more pieces of the early COVID-19 puzzle at a time when clarity was sorely needed — and became a vital tool in a rapidly evolving crisis.

Three years later, we are delighted to continue our partnership with UCSF as we enhance and update the original Atlas.



COVID-19 data can now be layered in the Health Atlas for consideration alongside other variables and demographics. This image maps the incidence of COVID-19 relative to asthma ED admissions.

ACT 2

Self-Reporting: The Challenge of Sparse & Sporadic Data

The very nature of the COVID-19 pandemic presented a significant challenge for information designers to overcome.

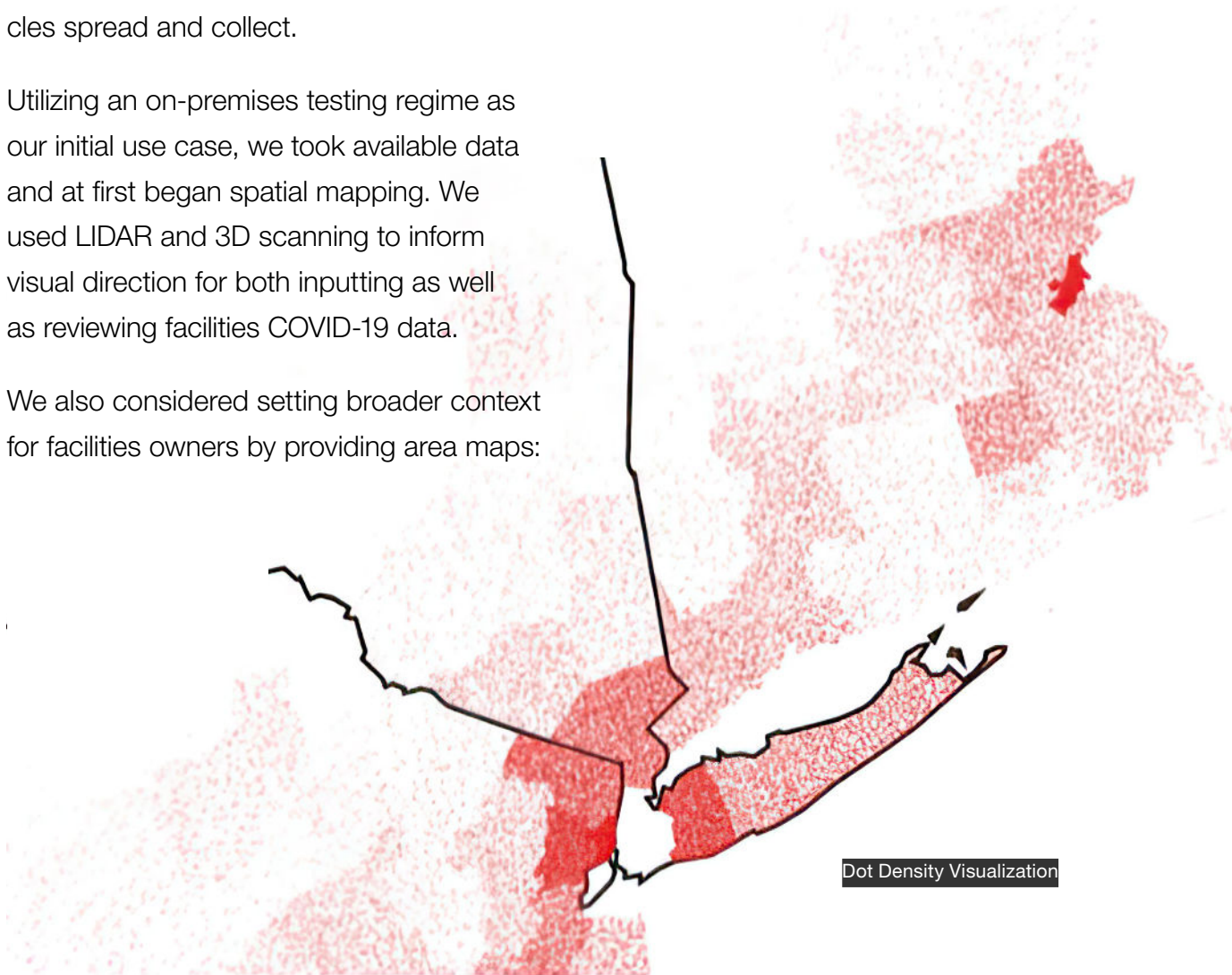
The COVID-19 drama played out across the globe and unfolded daily. Numbers changed rapidly. Organizations and corporations alike experimented with ways to capture the ever-changing flow of data. At Stamen, we worked on another COVID-19 project that highlighted unique challenges in gathering and visualizing pandemic data collected at the individual reporting level.

Phylagen: Tracking Coronavirus Within a Microbiome

Stamen solved data collection challenges for Phylagen, a biotech company that wanted to explore ways to track how coronavirus spread within the microbiome of office spaces. This project began by producing many interior space visualizations to further understand the way that microbiome particles spread and collect.

Utilizing an on-premises testing regime as our initial use case, we took available data and at first began spatial mapping. We used LIDAR and 3D scanning to inform visual direction for both inputting as well as reviewing facilities COVID-19 data.

We also considered setting broader context for facilities owners by providing area maps:



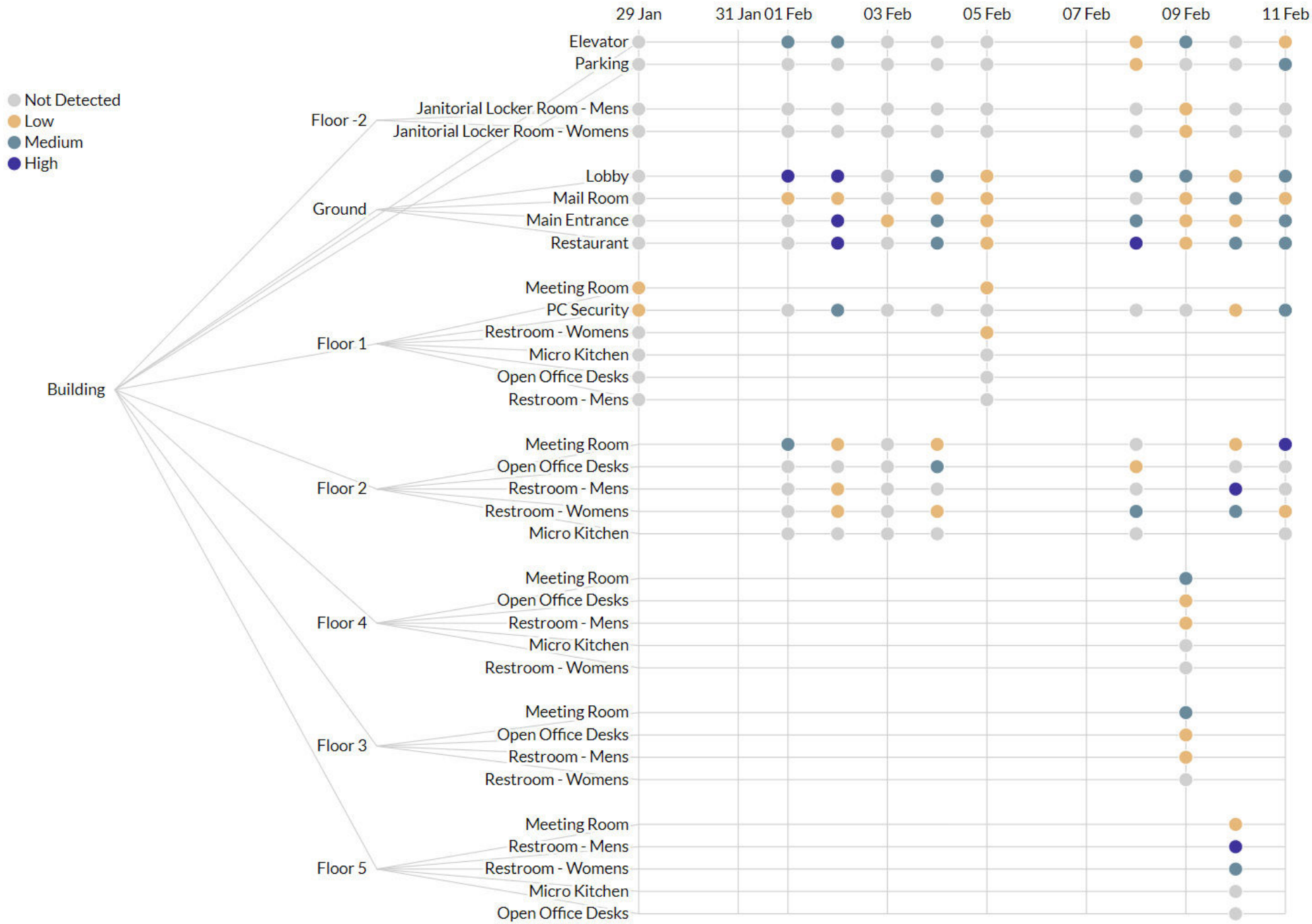
Dot Density Visualization

In the prototype for the use case, the user would then be able to zoom into an individual building and assess that by floor, as well as annotate the view for further discussion (image, right).

Outside of a few tooling use cases however, we were unable to make any visualizations leveraging maps or floor plans really work owing to the sparsity of the data; the actual utility of a visual tool remained unclear.

Consequently, we pivoted to focus exclusively on the final reporting aspect of the work and none of the data-acquisition components. We prototyped some data breakdowns with tree diagrams (image, next page).

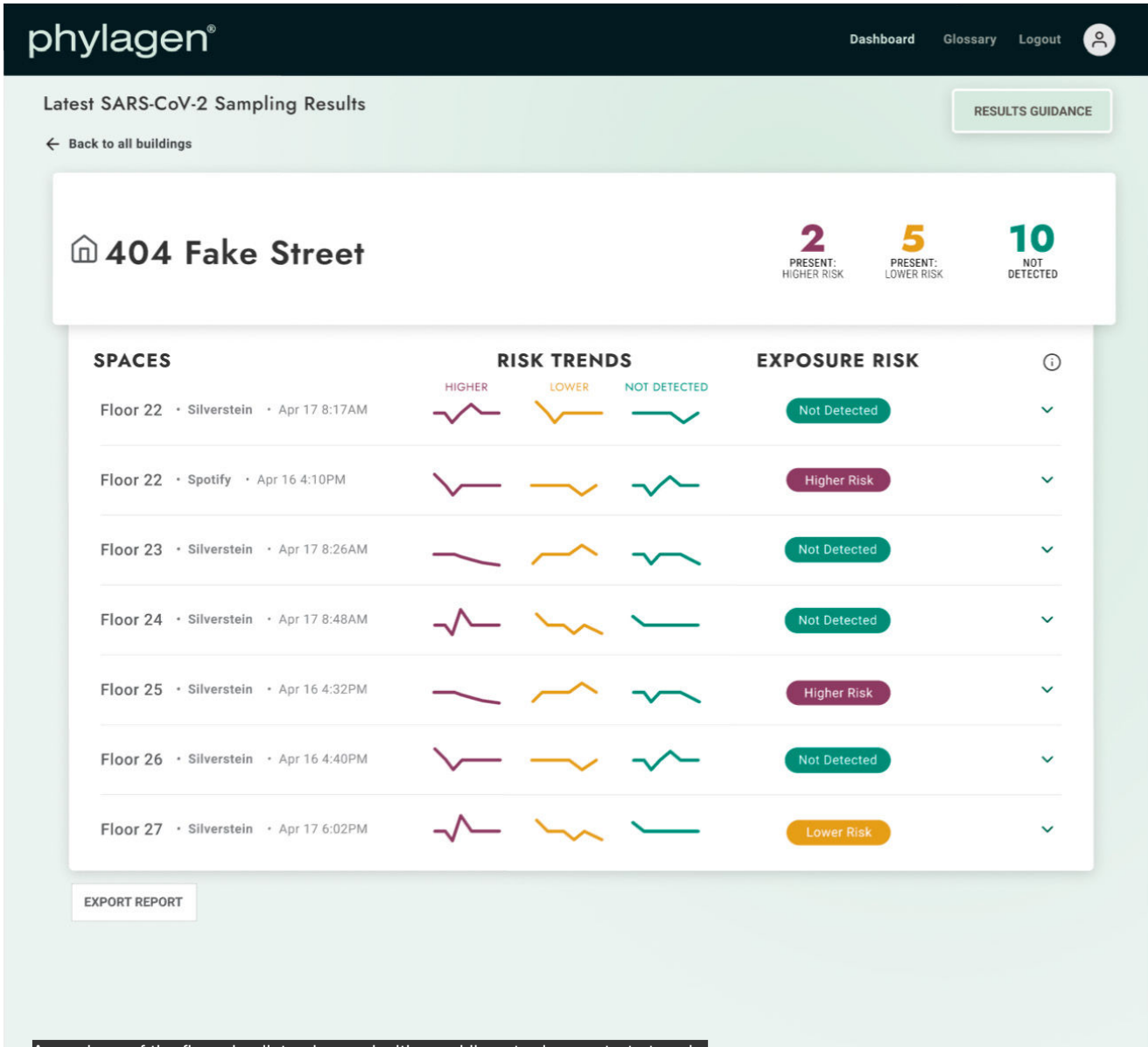
This example shows an early prototype that ingested building floor plans and then allowed for updating location data per room.


Not Detected
Low
Medium
High
29 Jan
31 Jan
01 Feb
03 Feb
05 Feb
07 Feb
09 Feb
11 Feb
Building
Floor -2
Janitorial Locker Room - Mens
Janitorial Locker Room - Womens
Ground
Elevator
Parking
Lobby
Mail Room
Main Entrance
Restaurant
Floor 1
Meeting Room
PC Security
Restroom - Womens
Micro Kitchen
Open Office Desks
Restroom - Mens
Floor 2
Meeting Room
Open Office Desks
Restroom - Mens
Restroom - Womens
Micro Kitchen
Floor 4
Meeting Room
Open Office Desks
Restroom - Mens
Micro Kitchen
Restroom - Womens
Floor 3
Meeting Room
Open Office Desks
Restroom - Mens
Restroom - Womens
Floor 5
Meeting Room
Restroom - Mens
Restroom - Womens
Micro Kitchen
Open Office Desks

But in the end an annotated tabular view seemed appropriate and sufficient (image, right).

We finished our engagement with Phylagen having passed through every Stamen discipline — cartography, data science, UX research and design, visual design, and product development — and ending up with a durable roadmap for creation of a reporting tool to satisfy their product goals.

Though COVID-19 presented a number of unique challenges to solve, the truth is that any data visualization project is experimental in nature. And since some ideas are better in theory than in application, arriving at the right solution for Phylagen required trial and error — iterating weekly to really clarify the user's needs and satisfy them. In the case of Phylagen, we started with maps & 3D telemetry and ended up with a fairly simple table. Phylagen quickly saw the benefits of working with Stamen's design &



A mock up of the floorplan list enhanced with sparklines to demonstrate trends.

development team and our ability to inform strategy and pivot our focus was invaluable to the effort. Trying different angles and approaches is the only way to find the best path to tell the right data story, or to determine if a data project is truly viable and meaningfully accurate.

Themes of Act 2's COVID-19 Data Projects

Act 2 reveals three themes: exploration, investigation, and strategic input.

Exploration

The Phylagen initiative began as an exploratory project. Phylagen, alongside many other organizations, was eager to leverage any idea, data set, or product functionality available to increase public health safety. They recognized the value that data exploration could offer toward validating or invalidating the usefulness of their ideas.

Data exists along a vast spectrum of abstraction. Thus, data exploration seeks to learn about the characteristics and any potential problems of the data set without the need to formulate assumptions beforehand.

Data exploration answers important questions:

- Is the available data suitable to achieve the project goals?
- Do we know enough about why any data gaps exist?
- Which data sets and data subsets can or cannot be used for building visual models?

Data exploration is an approach rapidly gaining popularity as organizations increasingly embrace data science and interactive visualizations. Data exploration is key to the discovery phase of any Stamen project because it informs an initiative with two key advantages:

1. Data exploration enables unexpected discoveries in the data.
2. Data exploration fosters a deep understanding of the data for project viability, efficiency, and success.

Because humans have powerful visual pattern recognition capabilities, data exploration is most effectively conducted through visualizations.

You can stare at a numerical report all day and never see what becomes quickly apparent when looking at those same numbers in a well-crafted visualization.

Visualizing Phylagen's data during the exploratory phase empowered the discovery of patterns that were not anticipated and would have been hard to understand without experimenting with visual storytelling forms.

Investigation

Investigating variables revealed during data exploration is the natural next step in a project's discovery process.

Investigation in the Phylagen project involved visualization prototypes and design mockups. Because COVID-19 data products are complex and nuanced, prototyping the visual interfaces of the Phylagen dashboards and data views provided experimental models for how their data could be rendered.

Strategic Input

Through iteration with Phylagen product, engineering and data science teams, Stamen provided each team with various tools to explore data views and wireframes. However, discovery efforts ultimately proved the logistical challenges of gathering actionable data in a timely way. The core product offering of the prototype application was thus fundamentally changed to a view-only focus on clear and simple broadcasting of available data.

ACT 3

Wastewater Data: Setting the Gold Standard

If you've been keeping an eye on the latest data regarding COVID-19 and other viruses, you've likely heard that much of the best data we have today comes from wastewater surveillance.

In the course of our work with the UCSF Health Atlas and Phylagen, we learned that when more testing began to happen at the individual level in people's homes, reported lab test data became less reliable. Data collection at the individual level was also problematic. It relied on individuals taking the time to upload or otherwise report their test results.

However, everyone who lives or works within a sewershed contributes waste to the treatment plant that services that area. And that waste contains a wealth of information about virus pathogens which are reliable and definitive.

Down the Sewers in Search of Data

WastewaterSCAN brings together researchers at Stanford and Emory Universities with industry and philanthropic groups to work with wastewater treatment plants. The goal is to monitor pathogen presence at no cost to participating municipalities. These groups seek to create data visualization and analysis tools that make it easier for public health departments to monitor health trends. Why? So officials can take necessary actions to protect the public as well as to inform and assure their constituents with stable, correct information.

Last year, the WastewaterSCAN team approached Stamen because they knew they needed to expand the reach and abilities of their project in two critical ways:

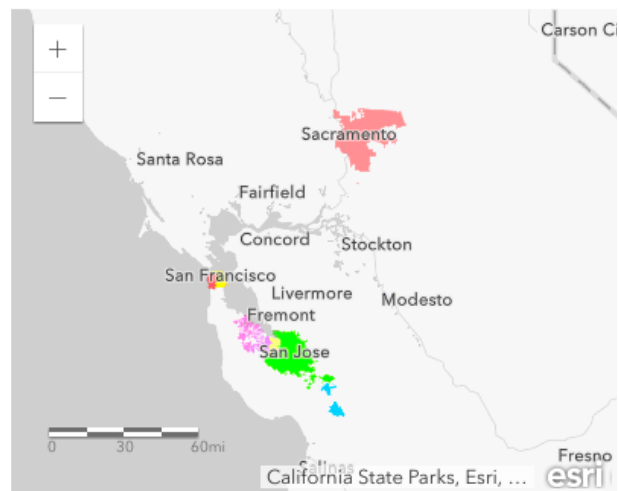
1. They wanted to test for more pathogens, not just COVID-19.
2. They wanted to collect data from an exponentially increasing number of wastewater treatment plants.

We recently published this article that digs into the details of WastewaterSCAN's project. What is compelling here is the contrast this project provides to the data projects mentioned in Act 1 and Act 2. The UCSF Health Atlas relied on data gathered from reporting sources like public newspapers or health reports. Phyalogen leveraged data gathered at individual levels. WastewaterSCAN captured the opportunity to passively and consistently collect data from an entire geographical area — ultimately providing the most accurate COVID-19 data collection methodology to date.

As in Act 1 and Act 2, the discovery process of this project proved critical to arriving at the right combination of visual forms to communicate WastewaterSCAN's data.

Plentiful Data & Actionable Deliverables

WastewaterSCAN already had some maps and charts that worked well for a set of plants reporting pathogen data from the Bay Area, but a more robust solution was needed.



One of WastewaterSCAN's early maps of sewershed areas from Bay Area wastewater treatment plants.

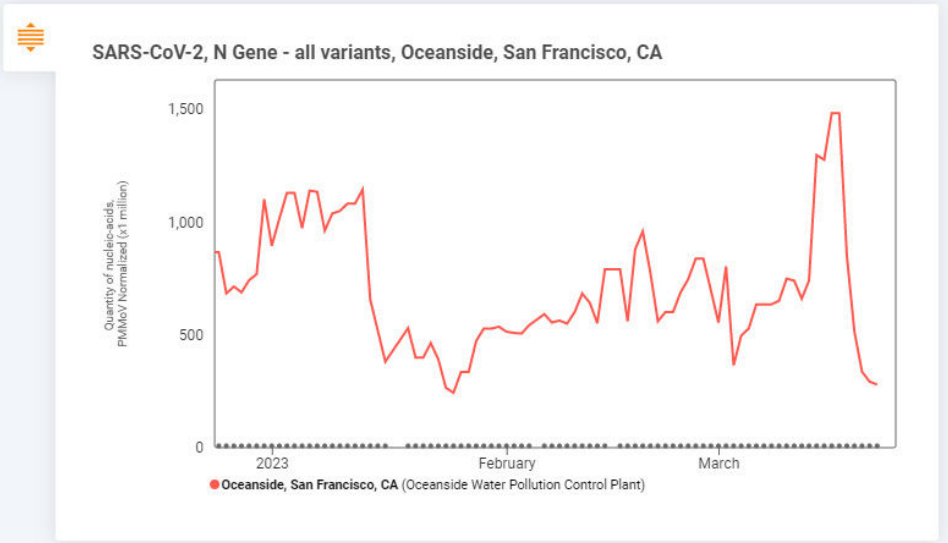
As is common on a project with a large amount of multi-faceted data, there were numerous ways to conceptualize the visualizations. We spent time working closely with the WastewaterSCAN team to aggregate multiple methods of visualization in a way that invites both advanced and novice users to glean insight.

We also needed to ensure accuracy. It was imperative to give researchers the right trendlines and heatmap charts that properly represented the data.

A chart builder (image, next page) offered the most compelling use case to allow users the ability to customize and export their analysis of WastewaterSCAN's findings.

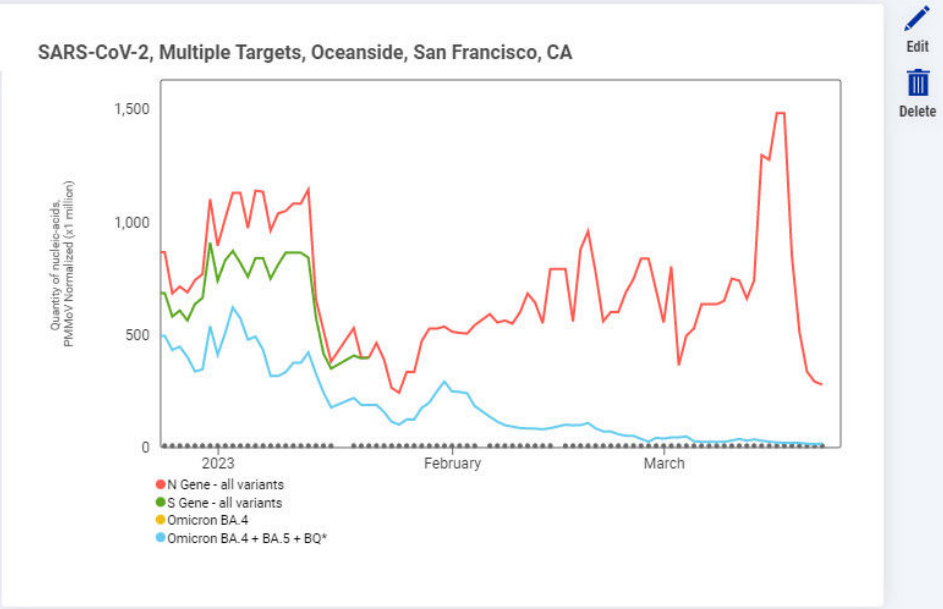
Sewershed Chart Grid Edit

<> Embed



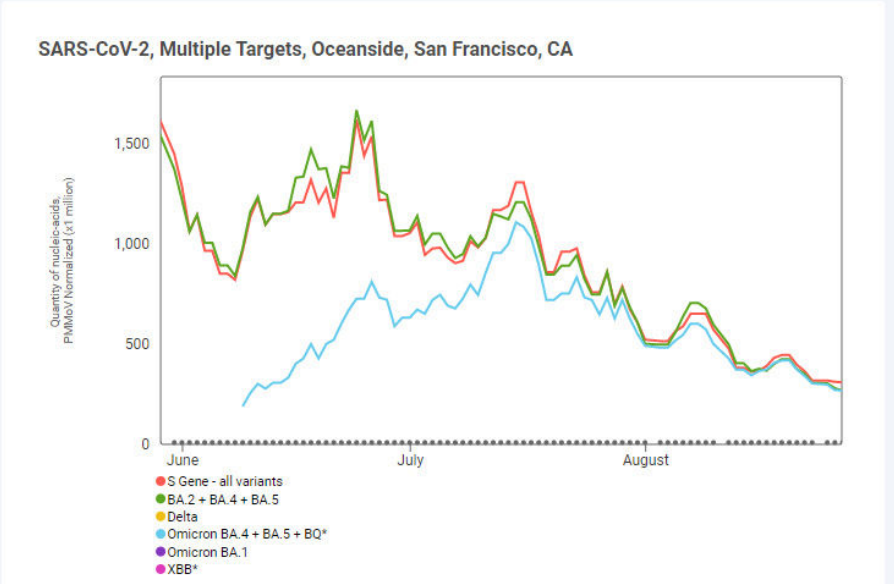
Edit

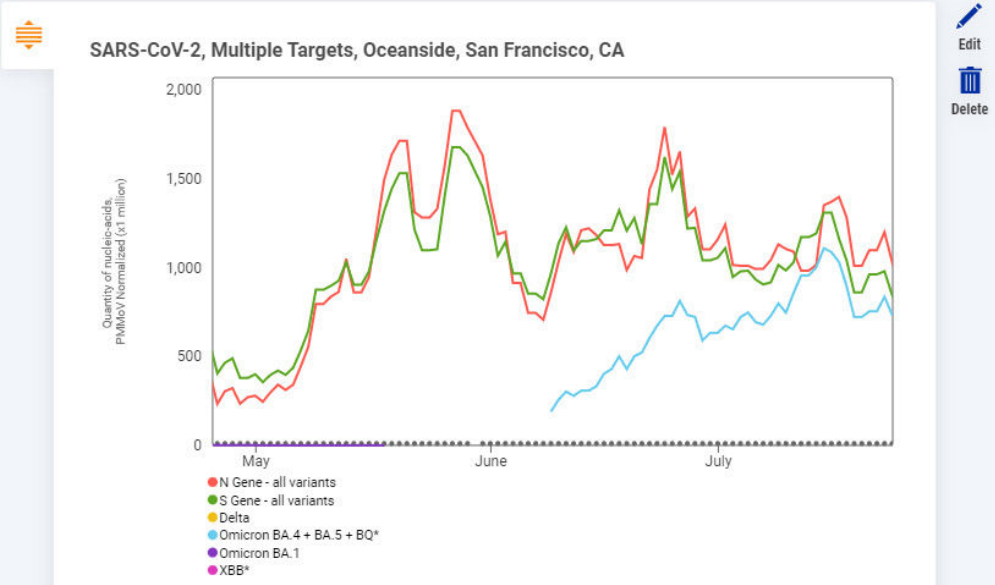
Delete



Edit

Delete





Edit

Delete

The ‘small-multiples’ view with several charts at once can also be embedded into websites for counties, cities and water treatment facilities to provide ongoing reporting on the health of their watersheds.

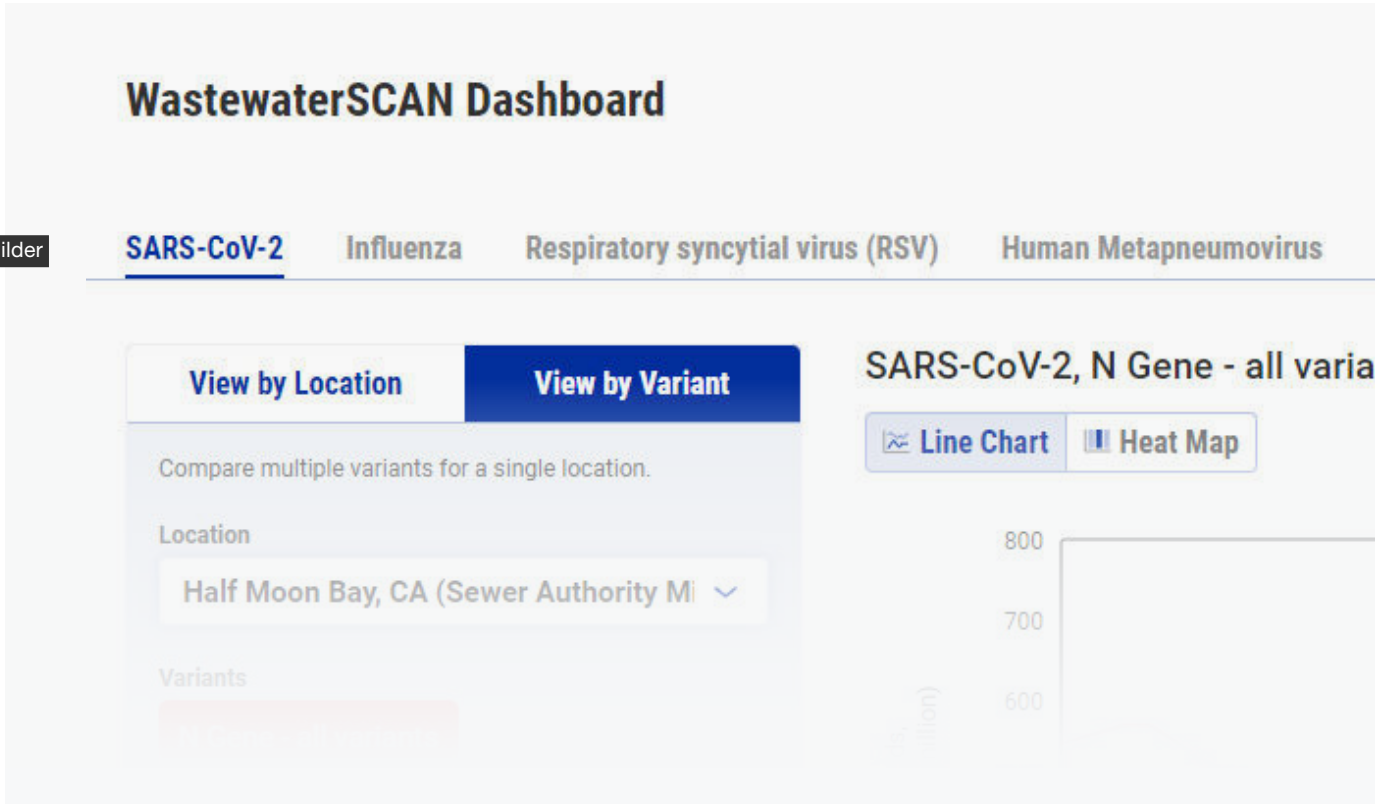
The ‘small-multiples’ view with several charts at once can also be embedded into websites for counties, cities and water treatment facilities to provide ongoing reporting on the health of their watersheds.

From a single view, users can view a combination of line charts and heatmaps to clearly identify trends and compare pathogen levels over time. They can choose the type of chart, which locations and pathogens to track, and a time range within which to view the data among a host of other exploration options.

Curating Choices to Provide More Value

In prototyping the chart builder, challenges surfaced regarding the ways users could compare data slices. The possibility of combining any slice of data with any other piece of data could cause accuracy issues or be potentially misleading. For example, directly comparing two different pathogens within the same chart could inadvertently confuse users because the trending quantity for different pathogens holds different meanings and different causes for concern.

Eliminating anything that might hamper a user’s experience in the chart builder required additional data hierarchy analysis and programming. For example, adding a tabbed pathogen selector emphasized the difference between pathogens.



Adding a tabbed pathogen selector in the chart builder emphasized the difference between pathogens.

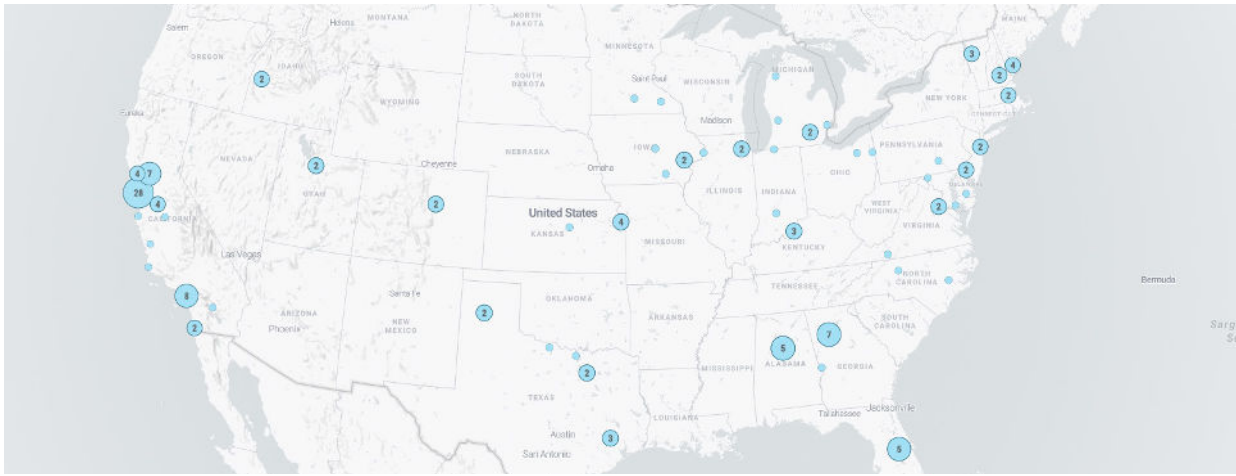
Curating a user experience is key to making intuitive visualizations

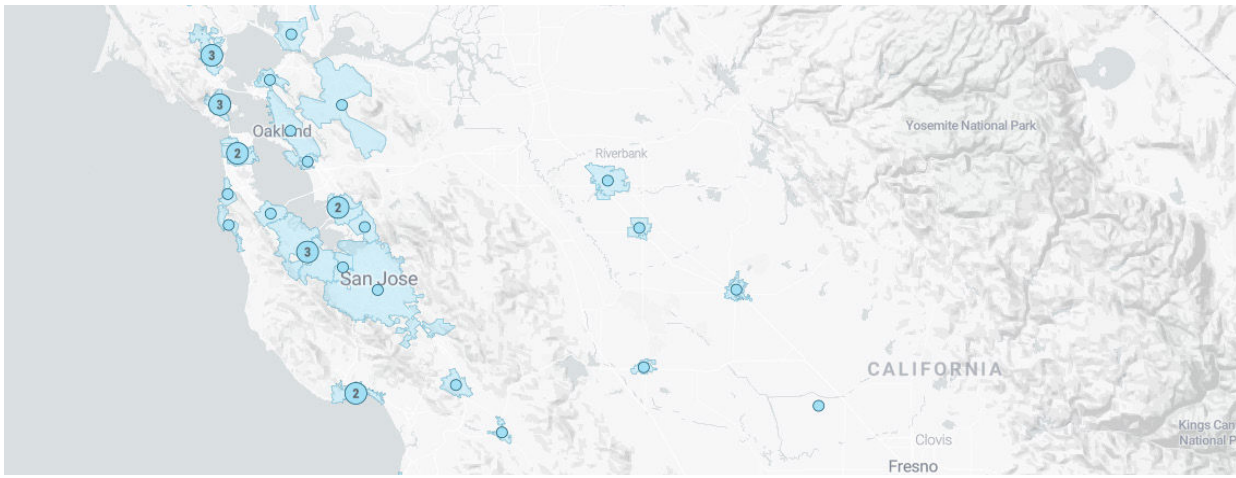
Users can select an area (top image) with participating wastewater treatment plants across the nation and view pathogens being reported at specific locations.

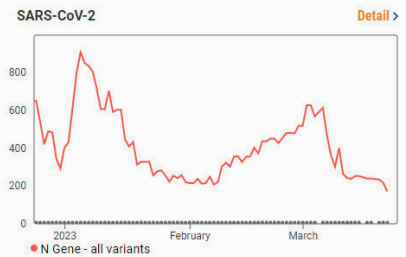
Users can also zoom in (middle image) to see the geographical boundaries of the sewershed that each plant treats.

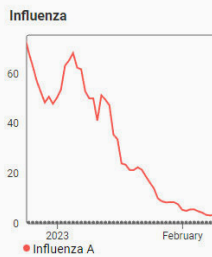
And, finally, users can select any location on the map (bottom image) to view a summary of any pathogen currently being reported.

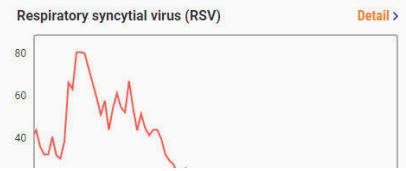
In future iterations, we hope to expand even more features that empower users to explore and share this data.

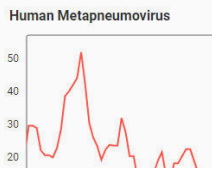










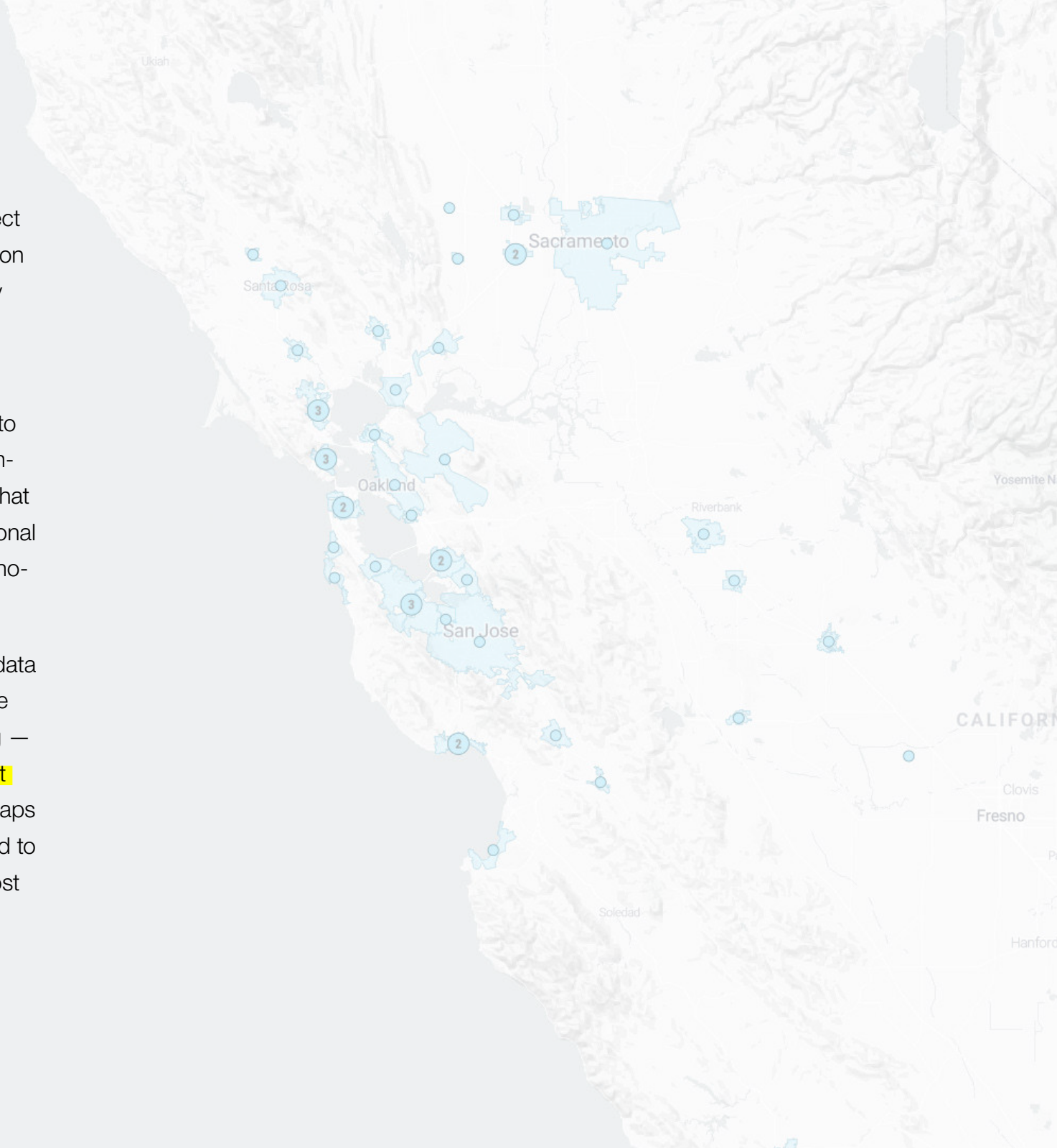


Themes of Act 3's COVID-19 Data Project

WastewaterSCAN's COVID-19 data project reinforces the critical value data exploration and visualization prototyping bring to any project.

In designing an end-to-end technical solution, we crafted an intuitive platform to ingest and manage large amounts of continually flowing data. This powers a tool that makes COVID-19 data available at a national scale in a normalized way to provide pathogen monitoring for the general public.

The creation of novel and sophisticated data products empowers people to have more informed conversations about everything — from COVID-19 to myriad other important topics. The development of interactive maps and data visualizations is uniquely tailored to each client, their data, and the data's most impactful use-cases.



Collaborating to Find a Solution

In many ways, the design & development team at Stamen had a front-row seat to the rapidly-changing landscape of COVID-19 data collection. Beginning with on-the-fly data collection methodologies, we generated meaningful visualizations to drive early decisions about public safety within the Bay Area. Today, we are working with the most sophisticated method of public data collection available for pathogen research and wastewater scans.

At Stamen, discovery is a comprehensive process encompassing end-to-end engineering solutions and intuitive visualization designs. Our designers meet you where you are at with your unique project, and together we explore all the possibilities of what your data-driven communication goals can ultimately become, and then jointly craft solutions that are durable, useful and beautiful.

Stamen

Helping people find and tell more beautiful stories with data

584 Castro Street #240
San Francisco, CA 94114

415.558.1610
info@stamen.com

37° 45’ 34” N
122° 26’ 6” W